5.1 The pigeonhole principle

**Proposition**

Let $k, r$ be positive integers.

Let $A_1, A_2, \ldots, A_n$ be finite non-empty sets that are pairwise disjoint. Assume that $|A_1 \cup A_2 \cup \ldots \cup A_n| > kr$.

Then there exist at least one index $i$ such that $|A_i| > r$.

**Proof**

By contradiction.

Assume that $|A_i| \leq r$ for each $i$. Then

$$|A_1 \cup A_2 \cup \ldots \cup A_n| = |A_1| + \ldots + |A_n| \leq nr,$$

contradicts the assumption.
Remark

Which assumptions can be dropped? And the sets are allowed to be empty.
If \( r = 0 \) the claim still holds. In this case the principle asserts that if the union of the sets \( A_1, \ldots, A_k \) is non-empty, then at least one of the sets \( A_i \) has to be non-empty.

If the sets are not necessarily pairwise disjoint, the claim is still valid. Note that for two arbitrary finite sets \( A, B \) we have \( |A \cup B| \leq |A| + |B| \), for if \( x \in A \cup B \), then \( x \) is counted once on the left-hand side of the inequality, while it is counted twice on the right-hand side. (We will give a rigorous proof of this in Week 3.)

Thus we have
\[
|A_1 \cup \cdots \cup A_k| \leq |A_1| + \cdots + |A_k| \leq kr.
\]

Example

There are at least 120 students at UCLA celebrating their birthday on the same day of the year. Let \( A_i \) be the set of UCLA students born on the \( i \)th day of the year, \( i = 1, \ldots, 366 \).

We have
\[
|A_1 \cup \cdots \cup A_{366}| < 120 \times 366 = 43,560,
\]
and hence at least one of the sets \( A_i \) must contain more than 119 students.
5.2 Multisets

**Definition**

A multiset is a tuple \((S, m)\) where:
- \(S\) is a non-empty set
- \(m: S \to \mathbb{Z}_{\geq 0}\) is a map.

For each \(s \in S\) we call \(m(s)\) the multiplicity of \(s\).
Let \(n = \sum_{s \in S} m(s)\). Then one says that \((S, m)\) is an \(n\)-element multiset over \(S\).

**Example**

The prime factors of a number, e.g., \(168 = 2^3 \cdot 3 \cdot 7\), give the multiset \(\{2, 2, 2, 3, 7\}\).

Note: by abuse of notation one sometimes also writes \(\{2, 2, 2, 3, 7\}\) as \(\{2, 2, 2, 3, 7\}\) by listing the elements exactly each element \(s\) of \(S\) \(m(s)\) many times, so in the example above we would have \(\{2, 2, 2, 3, 7\}\).
5.2.1 Weak compositions
Given an \( n \)-element multiset over \([n]\), we can associate to it the ordered tuple \((m(1), \ldots, m(n))\) given by the multiplicity of the elements of \([n]\).

Conversely, given an ordered tuple \((a_1, \ldots, a_k)\) of non-negative integers \(a_1, \ldots, a_k\) such that \(\sum_{i=1}^{k} a_i = n\), we can associate to such a tuple an \(n\)-element multiset over \([n]\). This assignment is a bijection.

**Defn**

Let \(a_1, a_2, \ldots, a_k\) be non-negative integers such that \(\sum_{i=1}^{k} a_i = n\). The ordered \(k\)-tuple \((a_1, \ldots, a_k)\) is called a **weak composition** of \(n\) into \(k\) parts.

Thus, counting the number of \(n\)-element multisets over \([n]\) is the same thing as counting the number of weak compositions of \(n\) into \(k\) parts.

5.2.2. The number of \(n\)-element multisets over \([n]\).

**Proposition**

Let \(n\) be a non-negative integer, \(k\) a positive integer.

Then the number of \(n\)-element multisets over \([n]\) is
Proof

Let \( ([n], m) \) be an \( n \)-element multiset over \([n]\). We can assign to it an ordered \( n \)-tuple \((x_1, \ldots, x_n)\) such that \( x_i \in [n] \) and \( 1 \leq x_1 \leq \ldots \leq x_n \leq n \), by listing the elements of \([n]\) as many times as given by their multiplicity. This assignment is a bijection. Thus, counting the number of \( n \)-element multisets over \([n]\) amounts to computing the cardinality of the set

\[ A := \{ (x_1, \ldots, x_n) \mid x_i \in [n] \text{ and } 1 \leq x_1 \leq \ldots \leq x_n \leq n \}. \]

We will compute the cardinality of \( A \) by constructing a bijection between \( A \) and a set for which it is easier to compute the cardinality.

Let \( S \) be an \( n \)-element subset of \([n+1]\). We can assign to \( S \) an ordered \( n \)-tuple \((\beta_1, \ldots, \beta_n)\) such that \( \beta_i \in [n+1] \) and \( 1 \leq \beta_1 < \ldots < \beta_n \leq n+1 \), by listing the elements of \( S \) in increasing order.
Let \( B := \{ (\beta_1, \ldots, \beta_n) \mid \beta_i \in [n+k-1] \text{ and } 1 \leq \beta_1 < \beta_2 < \cdots < \beta_n \} \).

**Notation:** \((x_1, \ldots, x_n) := (x_1, \ldots, x_n) \).

Define \( f : B \rightarrow A \)

\[(\beta_i)_{1 \leq i \leq n} \mapsto (\beta_i - i + 1)_{1 \leq i \leq n}.\]

This is a bijection, so the number of \( n \)-element multisets over \([n]\) is \( \binom{n+k-1}{n} \). \( \square \)

**Remark**

In a multiset \((S, m)\) we are allowing \( m(s) = 0 \) for some \( s \in S \), so we are allowing for elements of \( S \) to have multiplicity zero; e.g. \([3, 11, 3, 4, 4, 4, 3]\) is a 6-element multiset of \([5]\).